INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Series Expansions of Lyapunov Exponents and Forgetful Monoids

Stéphane Gaubert — Dohy Hong

N° 3971

July, 2000

THÈMES 1 et 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--003971--FR

Series Expansions of Lyapunov Exponents and Forgetful Monoids

Stéphane Gaubert* , Dohy Hong†

Thèmes 1 et 4 — Réseaux et systèmes — Simulation et optimisation
de systèmes complexes
Projets MCR et METALAU

Rapport de recherche n° 3971 — July, 2000 — 27 pages

Abstract: We consider Lyapunov exponents of random iterates of monotone homogeneous maps. We assume that the images of some iterates are lines, with positive probability. Using this memory-loss property which holds generically for random products of matrices over the max-plus semiring, and in particular, for Tetris-like heaps of pieces models, we give a series expansion formula for the Lyapunov exponent, as a function of the probability law. In the case of rational probability laws, we show that the Lyapunov exponent is an analytic function of the parameters of the law, in a domain that contains the absolute convergence domain of a partition function associated to a special "forgetful" monoid, defined by generators and relations.

Key-words: Lyapunov exponents, random products, nonexpansive maps, automata, prefix codes, perturbation theory. *2000 Mathematics Subject Classification*: Primary 37M25; Secondary 37H15, 68Q70, 74H10.

This work was partially supported by the European Community Framework IV Program through the Research Network "ALAPEDES" (The Algebraic Approach to Performance Evaluation of Discrete Event Systems, RB-FMRX-CT-96-0074).

* ENSTA and INRIA. *Postal address*: UMA, ENSTA, 32 Bd. Victor, 75739 Paris Cedex 15, France. Fax: +33 1 45525282. E-mail: Stephane.Gaubert@inria.fr

† ENS, LIENS, 45 rue d'Ulm, 75230 Paris Cedex 05, France. Fax: +33 1 44322151. E-mail: Dohy.Hong@ens.fr

Développement en Série des Exposants de Lyapounov et Monoïdes Oublieux

Résumé : Nous étudions les exposants de Lyapounov des itérées de fonctions monotones homogènes. Nous supposons que les images de certaines itérées sont des droites avec probabilité positive. En utilisant cette propriété de perte de mémoire qui est vérifiée pour les produits aléatoires de matrices dans le semi-anneau max-plus, et en particulier, pour les modèles de tas de pièces du type Tetris, nous donnons une formule de développement en série pour l'exposant de Lyapounov en fonction de la loi de probabilité. Dans le cas d'une loi de probabilité rationnelle, nous montrons que l'exposant de Lyapounov est une fonction analytique des paramètres de la loi, dans un domaine qui contient le domaine de convergence absolue d'une fonction de type partition associée à un monoïde particulier dit "oublieux", défini par générateurs et relations.

Mots-clés : exposants de Lyapounov, produits aléatoires, fonctions contractantes, automate, codes préfixes, théorie des perturbations *Classification AMS 2000*: Primaire 37M25; Secondaire 37H15, 68Q70, 74H10.

Contents

1 Introduction **3**

2 Probability Measures on Words and Lyapunov Exponents **6**
2.1 Notation and Definitions 6
2.2 Cesaro Sum and Abelian Representations 8
2.3 Rational Probability Measures 10
2.4 Furstenberg's Cocycle Formula 10

3 Forgetful Monoids **12**
3.1 Presenting Forgetful Monoids 12
3.2 Random Walks in Forgetful Monoids 13
3.3 Partition-like functions 14

4 Series Expansions of Lyapunov Exponents **17**
4.1 Case of Stationary Probability Measures 17
4.2 Case of Nonstationary Rational Probability Measures 19

5 Applications **21**
5.1 The Case $\mathscr{F} = \{a^c\}$ under a Bernoulli Measure 21
5.2 Random Heaps of Pieces 23
5.3 Multiple Memory Loss Relations 24

1 Introduction

We say that a map $f : \mathbb{R}^d \to \mathbb{R}^k$ is *monotone* if for all $x, y \in \mathbb{R}^d$, $x \le y \implies f(x) \le f(y)$, where $\le$ denotes the usual product ordering of $\mathbb{R}^n$, for all n. We say that f is (additively) *homogeneous* if for all $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^d$, $f(\lambda + x) = \lambda + f(x)$, where for all $y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, $\lambda + y$ denotes the vector with entries $\lambda + y_i$. We will use the notation $\mathsf{t}y = \max_{1\le i\le n} y_i$ and $\mathsf{b}y = \min_{1\le i\le n} y_i$ (t and b stand for "top" and "bottom", respectively).

Given a stationary random sequence $f_1, f_2, \ldots$ of monotone homogeneous maps $\mathbb{R}^d \to \mathbb{R}^d$, and a vector $x \in \mathbb{R}^d$, we call *top Lyapunov exponent* the limit:

$$\Gamma = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathsf{t} f_n \circ \cdots \circ f_1(x)] \ . \tag{1}$$

Thus, the top Lyapunov exponent measures the linear growth rate of the orbits of the random dynamical system:

$$X_n = f_n(X_{n-1}) \qquad X_0 = x \ . \tag{2}$$

As observed by Vincent [47] (see §2 below for details), the limit in (1), if it exists, is independent of x, and, when $\mathsf{t} f_1(0)$ is integrable, the existence of the limit follows from the fact that the sequence $S_n = \mathbb{E}[\mathsf{t} f_n \circ \cdots \circ f_1(0)]$ is subadditive, i.e. $S_{n+k} \le S_n + S_k$. Moreover, if the sequence $f_1, f_2, \ldots$ is ergodic, Kingman's subadditive ergodic theorem shows that the Lyapunov exponent is also the almost sure limit:

$$\Gamma = \text{a.s.} \lim_{n\to\infty} \frac{1}{n}\mathsf{t} f_n \circ \cdots \circ f_1(x) \ . \tag{3}$$

A dual bottom Lyapunov exponent Γ' can be defined by replacing t by b in (1). Of course, all the results stated in this paper for Γ have dual versions for Γ'. More generally, we may replace t or b by an arbitrary monotone and homogeneous map $\varphi : \mathbb{R}^d \to \mathbb{R}$, and speak of φ-Lyapunov exponent. An interesting choice is the i-th coordinate map $\varphi(x) = x_i$, for some $1 \leq i \leq d$. Then, $\mathbb{E}\varphi(X_n) = \mathbb{E}(X_n)_i$ need not be subadditive or superadditive, and even in the case of a deterministic sequence ($f_1 = f_2 = \dots$), a counter-example, due to Gunawardena and Keane [29] (see Remark 1 below), shows that the φ-Lyapunov exponent, $\lim_n (X_n)_i/n$, need not exist (however, Theorem 2 below shows that under additional assumptions, the φ-Lyapunov exponent does exist).

The first example of monotone homogeneous map that we have in mind is

$$f : \mathbb{R}^d \to \mathbb{R}^d, \qquad f(x) = \log(M \exp(x)) \ , \tag{4}$$

where $\exp(x) = (\exp(x_1), \dots, \exp(x_d))^T$, $\log(x) = (\log(x_1), \dots, \log(x_d))^T$, and M is a $d \times d$ nonnegative matrix with at least one strictly positive entry per row (the later condition ensures that $f(\mathbb{R}^d) \subset \mathbb{R}^d$). If each map f_k is of the form $f_k(x) = \log(M_k \exp(x))$, for some M_k, then the Lyapunov exponent (1) coincides with the classical top Lyapunov exponent [11] of the random product of nonnegative matrices $M_n \dots M_1$, which is defined by:

$$\Gamma = \text{a. s.} \lim_{n\to\infty} \frac{1}{n} \log \|M_n \cdots M_1\| \ ,$$

for any norm $\|\cdot\|$.

The second and main example of monotone homogeneous map of interest to us is

$$f : \mathbb{R}^d \to \mathbb{R}^d, \qquad f_i(x) = \max_{1\leq j\leq d} (M_{ij} + x_j) \ , \tag{5}$$

where M is a $d \times d$ matrix with entries in $\mathbb{R} \cup \{-\infty\}$, such that each row contains at least one finite entry. If each map f_k is of the form (5) for some matrix M_k, the Lyapunov exponent (1) coincides with the Lyapunov exponent of the random product of matrices $M_n \dots M_1$ in the max-plus semiring [1, 15, 12].

An appealing example of max-plus random products is provided by Tetris-like heaps of pieces. For instance, consider the three monotone homogeneous maps $\mathbb{R}^2 \to \mathbb{R}^2$:

$$\begin{aligned} &a(x) = (x_1 + 1, x_2)^T, \qquad b(x) = (x_1, x_2 + 1)^T, \\ &c(x) = (\max(x_1, x_2) + 1, \max(x_1, x_2) + 1)^T, \end{aligned}$$

which are clearly of the form (5), the corresponding matrices being

$$A = \begin{pmatrix} 1 & -\infty \\ -\infty & 0 \end{pmatrix} \ , \quad B = \begin{pmatrix} 0 & -\infty \\ -\infty & 1 \end{pmatrix} \ , \quad C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \ .$$

To a sequence $f_1, \dots, f_n$, of elements of $\{a, b, c\}$, we associate the heap of pieces obtained by letting a sequence of n pieces with the corresponding shapes fall down on an horizontal ground. We denote by X_n the upper contour of the heap. For instance, Fig. 1 shows the heap of pieces corresponding to the sequence: a, b, a, c, b, b, b, a. Here, $n = 8$ and $X_n = (4, 6)$. It is quite easy to see that X_n is given by the random dynamical system (2) (see [24, 12, 26] for details). In this context, the Lyapunov exponent is equal to the almost sure limit of the height of a heap of pieces, divided by the number of pieces, when the number of pieces grows to infinity. When the sequence is i.i.d., and when the pieces a, b, c appear with the frequencies $p(a)$, $p(b)$, $p(c) = 1 - p(a) - p(b)$, respectively, the Lyapunov Exponent, which has been computed in [12], is the algebraic function represented in Fig. 2.

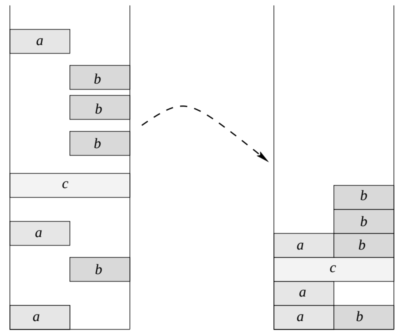
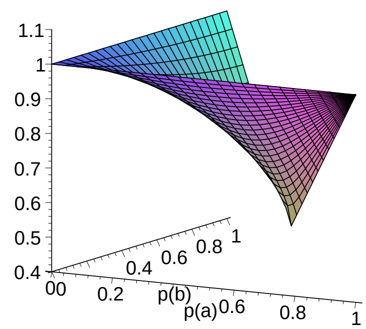


Figure 1: A two columns heap of pieces



$$\Gamma = \frac{1}{2}\,\frac{p(a)+p(b)-4p(a)p(b)+(2-p(a)-p(b))\sqrt{1-4p(a)p(b)}}{\sqrt{1-4p(a)p(b)}}.$$

Figure 2: Lyapunov exponent of the heap model of Fig. 1.

Computing exactly, or approximating, Lyapunov exponents of heaps of pieces, and more generally, of products of max-plus matrices, is a long standing problem [15, 44, 41, 1, 28, 23, 12, 3, 4, 25, 19, 10]. No exact formulæ are known, except in very special cases, such as the one of Fig. 2. In this paper, our purpose is rather to investigate the qualitative properties of Γ. For instance, a simple look at the formula in Fig. 2 shows that Γ is analytic in a domain which contains the set of real probabilities $\{0 \leq p(a), p(b) \leq 1, p(a)+p(b) < 1\}$, with a singularity at $p(a) = p(b) = \frac{1}{2}$. Hence, the power series expansion of Γ, seen as a function of $p(a), p(b)$, is convergent in any complex polydisc included in the domain $|p(a)| + |p(b)| < 1$. This is exactly what we prove here, in general: the main result of this paper (Theorem 2) shows that the Lyapunov exponent is analytic, and gives an explicit power series expansions, together with a tight estimate of its convergence domain. By summing this power series, we obtain a way to approximate Lyapunov exponents.

In general, the Lyapunov exponent need not be differentiable (look at the point $p(a) = p(b) = 1/2$ in Fig. 2), and it may even be discontinuous [43]. The critical assumption in our Theorem 2 is the *memory loss property*, whose importance, in the context of heaps of pieces, or more generally, of products of max-plus random matrices, has been recognized by several authors [37, 24, 12, 19]. For the heap model of Fig. 1, this assumption just means that the arrival of a rigid piece (piece c) occupying all the slots, resets the heap to a state identical, up to a vertical translation, to the initial state.

We give an analytic, elementary proof: we shall write the Lyapunov exponent as a Cesaro or Abel mean of a function on the free monoid, and, under the memory loss property, we shall see that this mean can be expressed as a sum over the elements of a monoid defined by generators and relations, that we call forgetful monoid. The relations are of the form $ua = u$, for all generators a, and for all u in a distinguished set of words $\mathscr{F}$ (the appearance of a factor u makes the product forget its right factor, which accounts for the name of the monoid). When the set $\mathscr{F}$ is rational, the associated forgetful monoid is nothing but a very special rational monoid [45, 42]. We obtain in passing an Abelian representation of the Lyapunov exponent (Theorem 1), which plays for Lyapunov exponents, mutatis mutandis, the role that resolvents plays for eigenvalues of linear maps. This representation is not used in the proof of Theorem 2, but we think that it illuminates the form of the series expansion, and that it is of interest per se.

The results of the present paper on the analyticity domain improve the ones proved previously by Baccelli and Hong: in [3], the analyticity domain was obtained from the explicit coefficients of Taylor series expansions, and in [4], it was obtained by a contraction argument for Hilbert's projective metric, inspired by Peres [43]. The proof technique that we use here is completely different: the explicit expansion formula that we obtain is much simpler to sum, its coefficients all are positive, and for this reason, we obtain a more accurate estimation of the analyticity domain. However, the memory loss property remains in essence, similar to the contraction properties in the projective space, used by Peres, and many others [11, 36, 32] (memory loss is indeed a very strong "ultimate" contraction property, of Lipschitz constant 0). It would be very interesting to prove similar results without contraction arguments. For instance, we do not know what our results become if one replaces the assumption "there exists an iterate that is strictly contracting for Hilbert's projective metric" by "the image of one iterate is a compact set in the projective space" (the Birkhoff-Hopf theorem [20] shows that both statements are equivalent in the special case of linear maps acting in the positive cone, hence, this question is only interesting for monotone homogeneous maps which are not linear in the usual sense, and in particular, for max-plus linear maps).

Let us conclude this introduction by mentioning some additional motivations, and related works. In the context of Discrete Event Systems, Lyapunov exponents measure the cycle time, i.e. the average time between two events (see [1] for an introduction). In Dynamic programming, the Lyapunov exponent measures the growth of the optimal cost or reward of deterministic optimal control problems, as a function of the horizon, when the transition costs or rewards are random [15]. Max-plus Lyapunov exponents also arise in Statistical Physics, in the study of disordered systems at low temperature [22, 17]. There is a number of contributions on Lyapunov exponents of products of random matrices. A classical one is the monograph [11]. A recent one, dedicated to the case of nonnegative matrices, is [32].

Some of the results of this paper have been announced in [2].

2 Probability Measures on Words and Lyapunov Exponents

2.1 Notation and Definitions

Given a finite alphabet Σ, we denote by Σ^k the set of words of length k, i.e. the set of sequences of the form $(a_1, \dots, a_k)$, with $a_1, \dots, a_k \in \Sigma$. The free semigroup on Σ is $\Sigma^+ = \cup_{k\geq 1}\Sigma^k$, equipped with the concatenation product: $(a_1, \dots, a_k)(b_1, \dots, b_l) = (a_1, \dots, a_k, b_1, \dots, b_l)$, for all $a_1, \dots, a_k, b_1, \dots, b_l \in \Sigma$. The free monoid Σ^* is obtained by adjoining to Σ^+ the empty sequence, which is called the empty word in this context.

Following Hansel and Perrin [30], we say that a map $p : \Sigma^* \to [0, 1]$ is a *probability measure* on words if $\sum_{a\in\Sigma} p(aw) = p(w)$, for all $w \in \Sigma^*$, and if $p(1) = 1$. This implies that $\sum_{w\in\Sigma^k} p(w) = 1$, for all k. We say that p is *stationary* if $p(w) = \sum_{a\in\Sigma} p(wa)$, for all $w \in \Sigma^*$. The term probability measure can be justified by associating with a word w the cylinder of base w, which is the set of left infinite words with suffix w: there is a unique probability measure on the σ-algebra generated by cylinders, such that the cylinder of base w has probability $p(w)$.

Let us now use this formalism when Σ is a finite set of monotone homogeneous maps $\mathbb{R}^d \to \mathbb{R}^d$. We shall adopt, throughout the paper, the following notation: if $g : \mathbb{R}^d \to \mathbb{R}^l$ and $f : \mathbb{R}^l \to \mathbb{R}^m$ are monotone homogeneous maps, and if $x \in \mathbb{R}^d$, we write gx for $g(x)$, fg for $f \circ g$, and we set

$$wx = a_k \circ \cdots \circ a_1(x) \quad \text{if} \quad w = (a_k, \dots, a_1) \in \Sigma^k \ . \tag{6}$$

Interpreting $p(w)$ as the probability that the first k elements of an infinite random sequence of elements of Σ are $a_1, \dots, a_k$, we rewrite the Lyapunov exponent (1) as:

$$\Gamma = \lim_{n\to\infty} \frac{1}{n} \sum_{w\in\Sigma^n} p(w)\varphi wx \ , \tag{7}$$

with the special choice $\varphi = \mathsf{t}$. The set $\mathbb{R}^d$, equipped with the action $\Sigma^* \times \mathbb{R}^d \to \mathbb{R}^d$, $(w, x) \mapsto wx$, might be thought of as a "non-linear automaton". Indeed, a conventional (deterministic) automaton is nothing but a finite state space equipped with an action of a free monoid, together with an initial state and a set of final states. Here, the state space is $\mathbb{R}^d$, the action involves monotone homogeneous maps, the initial state is $x \in \mathbb{R}^d$, and the set of final states is replaced by the output map φ. When passing from (1) to (7), we made the restriction that the set of maps Σ is finite. This is only to simplify the presentation, and to make clearer combinatorial aspects. The results of this paper have quite obvious extensions to the case of random iterates with continuous distributions.

To see when the limit (7) exists, it is useful to introduce:

$$S_n \overset{\text{def}}{=} \sum_{w\in\Sigma^n} p(w)\varphi wx = \mathbb{E}\varphi(X_n) \ . \tag{8}$$

We shall first consider the case when $X_0 = x = 0$. Since for all $w \in \Sigma^*$, the map $y \mapsto wy$, $\mathbb{R}^d \to \mathbb{R}^d$ is monotone and homogeneous, we have for all $u, v \in \Sigma^*$:

$$\mathsf{t}uv0 \le \mathsf{t}u0 + \mathsf{t}v0 \ , \tag{9}$$

for $uv0 \le u(0 + \mathsf{t}v0) = u0 + \mathsf{t}v0$. Hence, if p is stationary, the sequence S_n satisfies

$$\begin{aligned}
S_{n+k} &= \sum_{u\in\Sigma^n,\ v\in\Sigma^k} p(uv)\mathsf{t}uv0 \\
&\le \sum_{u\in\Sigma^n,\ v\in\Sigma^k} p(uv)(\mathsf{t}u0 + \mathsf{t}v0) \qquad \text{(by (9))}, \\
&= \sum_{u\in\Sigma^n} \Big(\sum_{v\in\Sigma^k} p(uv)\Big)\mathsf{t}u0 + \sum_{v\in\Sigma^k} \Big(\sum_{u\in\Sigma^n} p(uv)\Big)\mathsf{t}v0 \\
&= S_n + S_k \qquad \text{(by stationarity of } p).
\end{aligned}$$

By an elementary classical result, if a sequence S_n is subadditive (i.e. if $S_{n+k} \le S_n + S_k$), then the limit $\lim_n S_n/n$ exists. This shows that Γ exists when $x = 0$. To show that Γ exists for any x, it suffices to use the following classical easy observation [18]: a monotone homogeneous map f is non-expansive for the sup norm $\|\cdot\|$, i.e. $\|f(y) - f(y')\| \le \|y - y'\|$, for all $y, y' \in \mathbb{R}^d$. Hence, if $S_n(y)$ and $S_n(y')$ denote the sums (8) evaluated with $x = y$ and $x = y'$, respectively, we get $\|S_n(y) - S_n(y')\| \le \sum_{w\in\Sigma^n} p(w)\|\varphi wy - \varphi wy'\| \le (\sum_{w\in\Sigma^n} p(w))\|y - y'\| = \|y - y'\|$, which implies that $\lim_n S_n(y)/n = \lim_n S_n(y')/n$. We have reproved the following result, due to Vincent [47].

Proposition 1. *If p is a stationary probability measure and if $\varphi = \mathsf{t}$, then the Lyapunov exponent*

$$\Gamma = \lim_n S_n/n$$

exists, and is independent of $x \in \mathbb{R}^d$. □

Remark 1. If $\varphi \neq \mathsf{t}$, the Lyapunov exponent need not exist, as shown by the following variant, due to Sparrow [46], of the counter-example of Gunawardena and Keane [29]. Let $\Sigma = \{f\}$, with $f : \mathbb{R}^3 \to \mathbb{R}^3$, $(x_1, x_2, x_3)^T \mapsto (x_1, x_2 + 1, h(x_2 - x_1) + x_1)^T$, where h is any differentiable map such that $0 \leq h' \leq 1$. By construction, f is monotone and homogeneous. We get $f^k(0,0,0)^T = (0, k, h(k-1))^T$, for $k \geq 1$, and, setting $\varphi = x_3$, it is clear that we can choose h such that $\Gamma = \lim_k h(k)/k$ does not exist.

Remark 2. Of course, if the probability measure is not stationary, the Lyapunov exponent need not exist. E.g., consider a left infinite word w in two letters a, b, and let w_k denote the suffix of w composed of the k rightmost letters of w. Setting $p(z) = 1$ if z is a suffix of w, and $p(z) = 0$ otherwise, defines a probability measure on $\{a, b\}^*$. Now, take a, b to be the maps $\mathbb{R} \to \mathbb{R}$: $a(x) = x$, and $b(x) = x + 1$. We have $S_n = \mathsf{t} w_n 0 = |w_n|_b$, the number of occurrences of the letter b in w_n. It suffices to take an infinite word such that $\Gamma = \lim_n |w_n|_b/n$ does not exist to have the desired counter-example.

2.2 Cesaro Sum and Abelian Representations

The definition of the Lyapunov exponent $\Gamma = \lim_n S_n/n$ suggests to rewrite Γ as a Cesaro mean:

$$\Gamma = \lim_{n\to\infty} \frac{1}{n}(D_1 + \cdots + D_n) \ , \tag{10}$$

where

$$D_k \stackrel{\text{def}}{=} S_k - S_{k-1} = \sum_{w\in\Sigma^k} p(w)\varphi w x - \sum_{w\in\Sigma^{k-1}} p(w)\varphi w x \tag{11}$$

$$= \sum_{a\in\Sigma, w\in\Sigma^{k-1}} p(aw)\varphi a w x - \sum_{w\in\Sigma^{k-1}} p(w)\varphi w x \tag{12}$$

$$= \sum_{w\in\Sigma^{k-1}} D(w) \ , \tag{13}$$

with

$$D(w) = \sum_{a\in\Sigma} p(aw)(\varphi a w x - \varphi w x) \ .$$

Given a formal parameter q, we set

$$\Gamma_q = \sum_{w\in\Sigma^*} q^{|w|} D(w) \ , \tag{14}$$

where $|w|$ denotes the length of w. As is well known, in the case of a linear operator A, much information on the asymptotics of A^k when $k \to \infty$ can be derived by looking at the singularities of the resolvent $(\lambda - A)^{-1}$. As shown by the following theorem, a similar (but weaker) property holds for Lyapunov exponents, the role of the resolvent being played by Γ_q (therefore, we might call Γ_q the "Lyapunov resolvent").

Theorem 1 (Abelian Representation). *Let p denote a probability measure. If the Lyapunov exponent Γ, defined by the limit* (7)*, exists, then,*

$$\Gamma = \lim_{q\to 1^-} (1-q)\Gamma_q \quad . \tag{15}$$

Conversely, if $\varphi = \mathsf{t}$ or if p is stationary, and if $\lim_{q\to 1^-}(1-q)\Gamma_q$ exists, then Γ exists, and has the same value.

Proof. A well known result [31, Th. 55] states that if the Cesaro limit $\Gamma = \frac{1}{n}(D_1 + \cdots + D_n)$ exists, then, the Abelian limit $\lim_{q\to 1^-}(1-q)(qD_1 + q^2D_2 + \cdots) = \lim_{q\to 1^-}(1-q)\Gamma_q$ also exists and has the same value. This shows the first implication of the theorem. To show that the converse implication holds when $\varphi = \mathsf{t}$ or when p is stationary, we shall use the Tauberian theorem of Hardy and Littlewood [31, Th. 94], which states that if the limit $\lim_{q\to 1^-}(1-q)(qD_1 + q^2D_2 + \cdots)$ exists, then the Cesaro limit $\frac{1}{n}(D_1 + \cdots + D_n)$ exists and has the same value provided that the sequence D_k is bounded from above. It remains to check the later property when $\varphi = \mathsf{t}$, or when p is stationary.

We first assume that $\varphi = \mathsf{t}$, and we set

$$K = \max_{a\in\Sigma} \mathsf{t}a0 \quad .$$

Using successively the monotonicity and homogeneity of $y \mapsto ay$, $\mathbb{R}^d \to \mathbb{R}^d$, we get $\mathsf{t}awx \le \mathsf{t}a(0 + \mathsf{t}wx) = \mathsf{t}a0 + \mathsf{t}wx \le K + \mathsf{t}wx$. This implies that $D_k \le \sum_{w\in\Sigma^{k-1}} \sum_{a\in\Sigma} p(aw)K = \sum_{w\in\Sigma^k} p(w)K = K$.

Next, when p is stationary, we can write a sum dual to (12):

$$\begin{aligned} D_k &= \sum_{w\in\Sigma^{k-1}, a\in\Sigma} p(wa)\varphi wax - \sum_{w\in\Sigma^{k-1}} p(w)\varphi wx \\ &= \sum_{w\in\Sigma^{k-1}} \sum_{a\in\Sigma} p(wa)(\varphi wax - \varphi wx) \quad . \end{aligned} \tag{16}$$

Using the fact that for all $w \in \Sigma^*$, the map $y \mapsto \varphi wy$, $\mathbb{R}^d \to \mathbb{R}$, which is monotone and homogeneous, is non-expansive for the sup-norm, we get: $\|\varphi wax - \varphi wx\| \le \|ax - x\| \le K'$, where $K' = \max_{b\in\Sigma} \|bx - x\|$. Together with (16), this implies that $D_k \le K'$. $\square$

The following counter example shows that the converse implication in Theorem 1 need not hold when $\varphi \ne \mathsf{t}$ and when p is not stationary.

Example 1. Let $\Sigma = \{a, b, c\}$, where a, b, c are the maps $\mathbb{R}^3 \to \mathbb{R}^3$ such that

$$ax = (x_1 + 1, x_2 - 1, x_3)^T, \quad bx = (x_1, x_2, x_1)^T, \quad cx = (x_1, x_2, x_2)^T \quad .$$

We take $\varphi x = x_3$. Let w denote the periodic left infinite word $\ldots cabacaba$. As in Remark 2, we associate to w a probability measure. Here, the sequence $(S_k)_{k\ge 1}$ is equal to $(0, 1, 1, -2, -2, 3, 3, -4, -4, 5, 5, \ldots)$, $\Gamma = \lim_{k\to\infty} S_k/k$ does not exist. However, $\lim_{q\to 1^-}(1-q)\Gamma_q$ does exist. To see this, let us recall that a sequence $(s_k)_{k\ge 0}$ is m-Cesaro summable to ℓ if, defining inductively $S_k^0 = s_k$, and $S_k^r = S_1^{r-1} + \cdots + S_k^{r-1}$ for all $r \ge 1$, we have: $\lim_{k\to\infty} m! k^{-m} S_k^m = \ell$. Applying this definition to the sequence $s_k = D_k$, given by (11), we get $S_k^1 = S_k$, and it is easy to see that $(D_k)_{k\ge 0}$ is 2-Cesaro summable to 0. Since for any m, m-Cesaro summability implies Abel summability [31, Th. 43], we get that $(D_k)_{k\ge 0}$ is Abel summable to 0, which means that $\lim_{q\to 1^-}(1-q)\Gamma_q = 0$.

2.3 Rational Probability Measures

An interesting special case arises when the probability measure p is parametrized by finitely many coefficients. Let $\mathbb{R}_+$ denote the set of nonnegative reals. We say that a probability measure p on Σ^* is *rational* [30] if there exists an integer r, a row vector $\alpha \in \mathbb{R}_+^{1\times r}$, a column vector $\beta \in \mathbb{R}_+^{r\times 1}$, and a morphism $P : \Sigma^* \to \mathbb{R}_+^{r\times r}$, such that $p(w) = \alpha P(w)\beta$. We say that (α, P, β) is a (nonnegative) *linear representation* of dimension r of P. We will extend these notations to complex valued p, α, P, β, even if it has no probabilistic interpretation.

As observed in [30], a Bernoulli probability measure, which is of the form $p(a_k \dots a_1) = p(a_k)\cdots p(a_1)$ for all $a_1, \dots, a_k \in \Sigma$, is trivially rational, since it has the linear representation of dimension 1: $(1, p, 1)$. Markov measures, which are defined by $p(a_k \dots a_1) = \mathscr{P}(a_k, a_{k-1}) \dots \mathscr{P}(a_2, a_1)\pi(a_1)$, for some $\Sigma\times\Sigma$ column stochastic matrix $\mathscr{P}$ and for some stochastic vector π, are rational. Indeed, setting $\beta_a = \pi(a)$, $P(a)_{cb} = \mathscr{P}(c, b)$ if $a = b$, and $P(a)_{cb} = 0$ otherwise, and $\alpha_a = 1$ for all $a \in \Sigma$, it is easy to check that $p(w) = \alpha P(w)\beta$ and $p(w) = \sum_{a\in\Sigma} p(wa)$. If π is an invariant measure of P, then, p is stationary. Our definition of Markov measures coincides with that [43], except that we do not require the stationarity.

In the case of rational probability measures, the representations of the Lyapunov exponent can be made more explicit. To each $y \in \mathbb{R}^d$, we associate the following $r \times r$ matrix:

$$\delta(y) = \sum_{a\in\Sigma} P(a)(\varphi a y - \varphi y) \ . \tag{17}$$

In the Bernoulli case, $\delta(y)$ is a scalar, which can be interpreted as the mean one step increment:

$$\delta(y) = \mathbb{E}(\varphi(X_1) - \varphi(X_0) | X_0 = y) \ .$$

An application of (10) and (13) yields

$$\Gamma = \lim_{n\to\infty} \frac{1}{n} \sum_{|w|\le n-1} \alpha\delta(wx)P(w)\beta \ . \tag{18}$$

Formula (18) shows that, in essence, the Lyapunov exponent is a mean of the increment function $\delta(y)$, taken on the set of states reachable from x by the action of Σ^*.

2.4 Furstenberg's Cocycle Formula

We next specialize to our discrete context Furstenberg's cocycle representation of the Lyapunov exponent [21]. Given an action of Σ^* on a denumerable set S, $\Sigma^* \times S \to S$, $(w, y) \to w \cdot y$, we say that a map $\kappa : \Sigma^* \times S \to \mathbb{R}$ is a *cocycle* if $\kappa(uv, y) = \kappa(u, v \cdot y) + \kappa(v, y)$ holds for all $y \in S$ and $u, v \in \Sigma^*$. We say that κ *represents* the map $\Sigma^* \to \mathbb{R}$, $w \mapsto \varphi w x$, if

$$\varphi w x = \varpi(w \cdot z_0) + \kappa(w, z_0) \ , \qquad \forall w \in \Sigma^* \tag{19}$$

for some $z_0 \in S$ (called *initial state*), and for some bounded map $\varpi : S \to \mathbb{R}$ (called *output function*). Associating to $a_1, \dots, a_k \in \Sigma$, the sequence $z_i = a_i \dots a_1 \cdot z_0$, $i = 1, \dots, k$, we rewrite (19) as:

$$\varphi a_k \dots a_1 x = \varpi(z_k) + \sum_{i=1}^{k} \kappa(a_i, z_{i-1}) \ . \tag{20}$$

(When S is finite, a cocycle representation is exactly a subsequential transducer [7] with output in the monoid $(\mathbb{R}, +)$.) If $a_1, a_2, \dots$ is a random sequence of independent identically distributed elements of Σ, taken with a Bernoulli law p, $z_1, z_2, \ldots$ is a denumerable Markov chain with values in S, and, if, for instance, this Markov chain is irreducible positive recurrent with invariant measure π, we get by applying to (20) the ergodic theorem:

$$\lim_k \frac{1}{k} \sum_{a_k,\ldots,a_1\in\Sigma} p(a_k \ldots a_1)\varphi a_k \ldots a_1 x = \sum_{z\in S, a\in\Sigma} p(a)\pi(z)\kappa(a,z) \enspace ,$$

as soon as the last sum is absolutely convergent. Then,

$$\Gamma = \sum_{z\in S, a\in\Sigma} p(a)\pi(z)\kappa(a,z) \enspace . \tag{21}$$

Furstenberg's choice of cocycle is essentially the following. We say that $x, y \in \mathbb{R}^d$ are *parallel*, and we write $x \parallel y$, if $x = \lambda + y$, for some $\lambda \in \mathbb{R}$. We call *line* generated by $x \in \mathbb{R}^d$ the equivalence class of x for the relation $\parallel$, namely $\mathbb{R} + x = \{\lambda + x \mid \lambda \in \mathbb{R}\}$. The (additive) *projective space*, $\mathbb{PR}^d$, is the set of lines. We take for S the set of lines of the form $\mathbb{R} + wx$, with $w \in \Sigma^*$. The homogeneity of φ and of the maps $z \mapsto wz$ allows us to equip S with the quotient action $\Sigma^* \times S \to S$, $(w, \mathbb{R}+y) \mapsto w \cdot (\mathbb{R}+y) \stackrel{\text{def}}{=} \mathbb{R} + wy$, and to take the cocycle $\kappa(w, \mathbb{R}+y) = \varphi wy - \varphi y$. Finally, taking the initial state $z_0 = \mathbb{R}+x$, and the constant output map $\varpi(\mathbb{R}+y) = \varphi x$, for all $\mathbb{R}+y \in S$, we obtain a cocycle representation of $w \mapsto \varphi wx$. The Lyapunov exponent of Fig. 2 has been computed using this technique. In general, the associated Markov chain z_k may not have recurrent states and we cannot apply formulæ like (21), but there are some important subcases where the analysis of z_k is simple. In particular, for max-plus linear maps with finite integer valued entries, the underlying Markov chain is finite [23], as in the following example.

Example 2. Consider the max-plus linear maps a, b associated respectively to the matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 2 & -1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 3 \\ 2 & 1 & 1 \end{pmatrix} \enspace .$$

We leave it to the reader to check that the image of a^2 is a line. Since Γ is independent of x, we can take as initial vector $x = a^2 0 = (2,3,3)^T$. The action of $\{a,b\}^*$ on x is depicted in Fig. 3. For instance, the arc marked $b, +5$ from $(6,6,5)^T$ to $(2,3,3)^T$ means that $b(6,6,5)^T =$

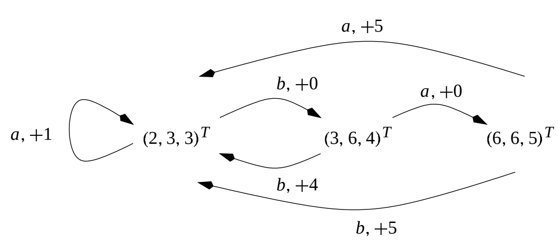


Figure 3: Cocycle Representation

$5 + (2,3,3)^T = (7,8,8)^T$. Thus, for any φ, we obtain a cocycle representation of $w \mapsto \varphi wx$ by setting $S = \{(2,3,3)^T, (3,6,4)^T, (6,6,5)^T\}$, by taking the initial state $z_0 = (2,3,3)^T$, the output map $\varpi(y) = \varphi y$, and, for $u \in \Sigma^*$ and $y \in S$, by taking for $u \cdot y$ the node reached from y by following the path with label u, and for $\kappa(u, y)$ the sum of the additive valuations on this path. The underlying Markov chain has transition matrix

$$\begin{pmatrix} 1-p(b) & p(b) & 0 \\ p(b) & 0 & 1-p(b) \\ 1 & 0 & 0 \end{pmatrix} ,$$

and it has the unique invariant measure

$$\pi = (1+2p(b)-p(b)^2)^{-1} \begin{pmatrix} 1, & p(b), & (1-p(b))p(b) \end{pmatrix} .$$

Applying (21), we get

$$\Gamma = \sum_{y\in S, c\in\Sigma} \pi(y)p(c)\kappa(y,c) = \frac{1+4p(b)-p(b)^2}{1+2p(b)-p(b)^2} . \tag{22}$$

3 Forgetful Monoids

3.1 Presenting Forgetful Monoids

The main ingredient in the explicit series expansion that we shall give in § 4, is an elementary monoid, that we next present.

Given a subset $\mathscr{F} \subset \Sigma^*$, we call *forgetful monoid* on Σ, with forgetful factors $\mathscr{F}$, the monoid with generators $a \in \Sigma$ and relations $wa = w$, for all $w \in \mathscr{F}$ and $a \in \Sigma$. We denote by $\mathbb{F}(\Sigma, \mathscr{F})$ this monoid. Formally, $\mathbb{F}(\Sigma, \mathscr{F})$ is the quotient of the monoid Σ^* by the least congruence $\equiv_{\mathscr{F}}$ such that $wa \equiv_{\mathscr{F}} w$, for all $w \in \mathscr{F}$ and $a \in \Sigma$ (congruences are identified to subsets of $\Sigma^* \times \Sigma^*$, ordered by inclusion). For instance, if $\Sigma = \{a, b, c\}$, and if $\mathscr{F} = \{c\}$, it is quite immediate to see that any word $w \in \Sigma^*$ is congruent for $\equiv_{\mathscr{F}}$ either to a word of the form zc, or to z, for some $z \in \{a, b\}^*$. This observation can be generalized, as follows.

We set:

$$\mathscr{S}_r = \Sigma^*\mathscr{F} - \Sigma^*\mathscr{F}\Sigma^+, \text{ and } \mathscr{S}_t = \Sigma^* - \Sigma^*\mathscr{F}\Sigma^* , \qquad \mathscr{S} = \mathscr{S}_r \cup \mathscr{S}_t , \tag{23}$$

where we use the standard notations for languages (if L, L' are languages, the concatenation LL' is the language whose elements are obtained by concatenating words in L with words in L'; the star L^* is defined by $L^* = L^0 \cup L \cup L^2 \cup \cdots$, where L^k denotes the k-th power of L for the concatenation product; $L^+ = LL^*$, and $L - L' = L \cap \complement L'$). It is quite easy to interpret $\mathscr{S}_r$ and $\mathscr{S}_t$: $\mathscr{S}_r$ is the set of words of the form uw, where $u \in \Sigma^*$ and $w \in \mathscr{F}$ is the only factor of uw belonging to $\mathscr{F}$, and $\mathscr{S}_t$ is the set of words that have no factors in $\mathscr{F}$ (recall that a word has a *factor* z if it can be written as uzv). The "r" and "t" in $\mathscr{S}_r$ and $\mathscr{S}_t$ stand for recurrent and transient, respectively (the terminology is justified by Prop. 3 below). Sets like $\mathscr{S}_r$ are known as *semaphore codes* [8, Chap. 2,§ 5].

Proposition 2. *We have the partition of Σ^* in equivalence classes modulo $\equiv_{\mathscr{F}}$:*

$$\Sigma^* = \bigcup_{z\in\mathscr{S}_t} \{z\} \cup \bigcup_{z\in\mathscr{S}_r} z\Sigma^* . \tag{24}$$

Thus, if $z \in \mathscr{S}_t$, the equivalence class of z is reduced to $\{z\}$, and if $z \in \mathscr{S}_r$, it is of the form $z\Sigma^*$.

Proof. Let $\mathscr{R}$ denote the relation such that $zwu\mathscr{R}zwv$ and $z\mathscr{R}z$, for all $z, u, v \in \Sigma^*$ and $w \in \mathscr{F}$. The relation $\mathscr{R}$ is reflexive and symmetric, by definition. Let us check that it is transitive. Let $t, t', t'' \in \Sigma^*$, such that $t\mathscr{R}t'\mathscr{R}t''$. If $t = t'$ or $t' = t''$, then $t\mathscr{R}t''$, trivially. Otherwise, we can write $t = zwu$, $t' = zwv = z'w'u'$ and $t'' = z'w'v''$ with $w, w' \in \mathscr{F}$ and $z, z', u, u', v, v'' \in \Sigma^*$. Since $zwv = z'w'u'$, either zw is a prefix of $z'w'$, or $z'w'$ is a prefix of zw. By symmetry, it is enough to consider the first case. Then, t, t', t'' all have zw as prefix, which implies that $t\mathscr{R}t''$. Thus, $\mathscr{R}$ is transitive. By definition, $z\mathscr{R}z' \implies zu\mathscr{R}z'u$ and $uz\mathscr{R}uz'$ for all $u, z, z' \in \Sigma^*$, which completes the proof that $\mathscr{R}$ is a congruence. Since $\mathscr{R}$ satisfies the presentation relations ($wa\mathscr{R}w$, for all $w \in \mathscr{F}$ and $a \in \Sigma$), $x \equiv_{\mathscr{F}} y \implies x\mathscr{R}y$. Conversely, $x\mathscr{R}y \implies x \equiv_{\mathscr{F}} y$, because for all $z, u, v \in \Sigma^*$ and $w \in \mathscr{F}$, the relations $zwu \equiv_{\mathscr{F}} zwv$ follow from $wa \equiv_{\mathscr{F}} w$ and from the fact that $\equiv_{\mathscr{F}}$ is a congruence. We have shown that the relations $\equiv_{\mathscr{F}}$ and $\mathscr{R}$ coincide. The partition (24) readily follows from the definitions of $\mathscr{R}$, $\mathscr{S}_t$, and $\mathscr{S}_r$. □

The interest in $\mathbb{F}(\Sigma, \mathscr{F})$ stems from the following observation, which is the key idea of the proof of Theorem 2 below. For all w, we denote by $\operatorname{Im} w$ the image of the map $\mathbb{R}^d \to \mathbb{R}^d,\ y \mapsto wy$.

Lemma 1. *If for all $w \in \mathscr{F}$, $\operatorname{Im} w$ is a line, we have, for all $z', z'' \in \Sigma^*$, and for all $x \in \mathbb{R}^d$:*

$$z' \equiv_{\mathscr{F}} z'' \implies z'x \parallel z''x \ . \tag{25}$$

Proof. Let us assume that $z' \equiv_{\mathscr{F}} z''$ and $z' \neq z''$ (otherwise, there is nothing to prove). By (24), z' and z'' are of the form zw' and zw'', for some $z \in \mathscr{S}_r$, $w', w'' \in \Sigma^*$. Then, $z'x$ and $z''x$, which belong to the same line, namely $\operatorname{Im} z$, are parallel. □

Thus, zx only depends of the equivalence class of z in $\mathbb{F}(\Sigma, \mathscr{F})$, up to an additive constant.

3.2 Random Walks in Forgetful Monoids

Given a sequence $u_1, u_2, \ldots$ of independent, identically distributed random variables with values in Σ, drawn with a Bernoulli distribution p, we define the left and right random walks X_k and Y_k on $\mathbb{F}(\Sigma, \mathscr{F})$, respectively, by $X_k = \overline{u}_k \cdots \overline{u}_1$ and $Y_k = \overline{u}_1 \cdots \overline{u}_k$, where $\overline{u}$ denotes the equivalence class of a word u modulo $\equiv_{\mathscr{F}}$.

The left random walk defines a denumerable Markov chain on $\mathbb{F}(\Sigma, \mathscr{F})$, to which we specialize the classical notions of accessibility, classes, recurrence, etc. In particular, when p is positive, we say that $X \in \mathbb{F}(\Sigma, \mathscr{F})$ *has access* to $X' \in \mathbb{F}(\Sigma, \mathscr{F})$ if there is a $Z \in \mathbb{F}(\Sigma, \mathscr{F})$ such that $X' = ZX$. A maximal set of mutually accessible elements is a *class*. A class whose elements only have access to elements of the same class is *final*. For the right random walk, these notions are defined in a dual way.

The following result is not needed in the subsequent proofs, but it shall give an intuitive interpretation to our main theorem (see Remark 4 below).

Proposition 3. *If p is a positive Bernoulli measure, then*

1. *$\{\overline{w} \mid w \in \mathscr{S}_r\}$ is the unique final class for the left random walk X_k;*
2. *The final classes for the right random walk Y_k are the one element sets $\{\overline{w}\}$, where $w \in \mathscr{S}_r$;*
3. *The unique invariant measure of the left random walk is: $\pi(\overline{w}) = p(w)$ if $w \in \mathscr{S}_r$, and $\pi(\overline{w}) = 0$ if $w \in \mathscr{S}_t$.*

Assertions 1 and 3 can be restated in a more appealing way as follows. If we write a word from right to left, drawing randomly each new letter of the word with the Bernoulli law p, and if, as soon as a prefix $f \in \mathscr{F}$ appears at the left of the word, we erase the part of the word at the right of f, we obtain a Markov chain with set of recurrent states $\mathscr{S}_r$, and the invariant measure, evaluated at a word $w \in \mathscr{S}_r$, is obtained by taking the product of probabilities of the letters of w.

Proof. Assertions 1 and 2 are clear. The restriction of the left random walk to its final class is clearly positive recurrent, hence, the invariant measure exists and is unique. To prove Assertion 3, it remains to show that π is normalized:

$$\sum_{w \in \mathscr{S}_r} p(w) = 1 \ , \tag{26}$$

and that it is invariant:

$$p(w) = \sum_{a \in \Sigma,\ z \in \mathscr{S}_r,\ az \equiv_{\mathscr{F}} w} p(a)p(z) \qquad \forall w \in \mathscr{S}_r \ . \tag{27}$$

We shall first derive (26) and (27) from results on codes, and then, we shall give a second, probably more intuitive, probabilistic proof.

Let us recall some definitions from [8] (the reader should consult this book for more details, and, in this proof, all references are relative to this source). A subset $X \subset \Sigma^*$ is a *code* if it generates a free monoid, and it is *prefix* if for any two words in X, none is a prefix of the other. Prefix sets, which are automatically codes, are called *prefix codes*. By construction, $\mathscr{S}_r$ is a prefix code. We say that a word $w \in \Sigma^*$ is *completable* in X if $uwv \in X$, for some $u, v \in \Sigma^*$. A code X is *thin* if there is one word not completable in X. For all $f \in \mathscr{F}$, ff is not completable in $\mathscr{S}_r$, thus, $\mathscr{S}_r$ is thin. Theorem 5.10 of Chap. 1 states that $\sum_{w \in X} p(w) = 1$ for all thin maximal codes X, and $\mathscr{S}_r$ is maximal by Cororollary 5.7 of Chap. 2. This shows (26). If $w = auf$, with $a \in \Sigma, u \in \Sigma^*$, $f \in \mathscr{F}$, (27) reduces to $p(w) = p(auf) = p(a)p(uf)$, which is true since p is Bernoulli. It remains to check (27) when $w = ag \in \mathscr{F}$, with $a \in \Sigma$. Then, (27) becomes

$$p(a)p(g) = \sum_{gu \in \mathscr{S}_r} p(a)p(g)p(u) \ ,$$

i.e., after canceling $p(a)p(g)$, $1 = \sum_{u \in g^{-1}\mathscr{S}_r} p(u)$, where $g^{-1}\mathscr{S}_r = \{w \in \Sigma^* \mid gw \in \mathscr{S}_r\}$. But Prop. 4.6 of Chap. 2 shows that $g^{-1}\mathscr{S}_r$ is a maximal prefix code, $g^{-1}\mathscr{S}_r$ is thin since $\mathscr{S}_r$ is thin, and Prop. 3.8 of Chap. 2, which states that $\sum_{w \in X} p(w) = 1$ for all thin maximal prefix codes X, yields (27).

Let us now give a probabilistic proof. Eqn (26) just says that an infinite word has a factor in $\mathscr{F}$ with probability one (this is an elementary fact that we shall not prove). To prove (27), we recall that $\pi(\overline{w})$ is equal to the mean frequency of visit of state $\overline{w}$ by the left random walk X_k. This frequency is the same if one considers the right random walk Y_k, because X_k and Y_k have the same distribution. Clearly, the frequency of visit of state $\overline{w}$ for the right random walk will be 1 if the sequence $u_1, u_2, \ldots$ begins by w, and 0 otherwise. Thus, the mean frequency is $\pi(\overline{w}) = p(w) \times 1 + (1 - p(w)) \times 0 = p(w)$. □

3.3 Partition-like functions

To a probability measure p, and a set $\mathscr{F} \subset \Sigma^*$, we associate the partition function:

$$Z = \sum_{w \in \mathscr{S}} p(w) \ . \tag{28}$$

We will sometimes write $Z(p)$, or $Z(\mathscr{F})$ to emphasize the dependence in p, or $\mathscr{F}$. We shall see in §4 that the convergence domain of the series expansion of the Lyapunov exponent is controlled by the convergence domain of $Z(p)$. In this preliminary section, we show how $Z(p)$ can be computed using standard methods from automata theory, when the probability measure p is rational, and when the set of forgetful factors $\mathscr{F}$ is a rational language.

First, we observe that when $\mathscr{F}$ is rational, $\mathscr{S}_r, \mathscr{S}_t$, and $\mathscr{S}$, all are rational (because rational languages are closed by product, star, and by the Boolean operations).

Next, to any language L, we associate the characteristic function $\operatorname{char} L : \Sigma^* \to \{0, 1\}$, which is defined by: $\operatorname{char} L(w) = 1$ if $w \in L$, and $\operatorname{char} L(w) = 0$, otherwise. Classically [9, Chap. 3, Prop. 1], a language L is rational iff its characteristic function is recognized by an automaton with multiplicities in the semiring of nonnegative integers $\mathbb{N}$, i.e. iff there exists an integer K, a row vector $I \in \mathbb{N}^{1\times K}$, a column vector $T \in \mathbb{N}^{K\times 1}$, and a morphism $\nu : \Sigma^* \to \mathbb{N}^{K\times K}$ such that $\operatorname{char} L(w) = I\nu(w)T$. In the sequel, we shall apply this construction when $L = \mathscr{S}$.

Last, we take a linear representation of p, (α, P, β). We denote by $\otimes$ the tensor or Kronecker product of matrices [38, Chap. 1,§ 1.9], and by $\nu \otimes P$ the map such that $(\nu \otimes P)(w) = \nu(w) \otimes P(w)$ for all $w \in \Sigma^*$.

We call *entries* of a linear representation (α, P, β) of dimension r all the terms of the form $\alpha_i, P(a)_{ij}, \beta_j$, where $1 \le i, j \le r$. If M is a matrix, we denote by $|M|$ the matrix with entries $|M_{ij}|$, and $\rho(M)$ denotes the spectral radius of M.

Proposition 4. *If $\mathscr{F}$ is a rational language, and if $p : \Sigma^* \to \mathbb{C}$ has the linear representation (α, P, β), Z is a rational function of the entries of (α, P, β), and the series (28) is absolutely convergent provided that*

$$\rho\Big(\sum_{a\in\Sigma} \nu(a) \otimes |P(a)|\Big) < 1 \ .$$

The proof of this proposition relies on a very classical identity, that we state as a lemma since it will be used several times in the sequel. If P is any morphism from Σ^* to a multiplicative monoid of square matrices with entries in $\mathbb{C}$, we define:

$$\hat{P} \stackrel{\text{def}}{=} \sum_{a\in\Sigma} P(a) \ . \tag{29}$$

Lemma 2. *If the spectral radius $\rho(\hat{P})$ is strictly less than 1, then, we have:*

$$\sum_{w\in\Sigma^*} P(w) = (1 - \hat{P})^{-1} \ . \tag{30}$$

Proof. We have $\sum_{w\in\Sigma^*} P(w) = \sum_{n\ge 0} \sum_{w\in\Sigma^n} P(w) = \sum_{n\ge 0} \hat{P}^n$. This series is absolutely convergent if $\rho(\hat{P}) < 1$. In this case, its value is $(1 - \hat{P})^{-1}$. □

We next prove Prop. 4. We have

$$\begin{aligned} Z &= \sum_{w\in\mathscr{S}} p(w) = \sum_{w\in\Sigma^*} \operatorname{char} \mathscr{S}(w) p(w) \\ &= \sum_{w\in\Sigma^*} I\nu(w)T\alpha P(w)\beta = \sum_{w\in\Sigma^*} (I \otimes \alpha)(\nu(w) \otimes P(w))(T \otimes \beta) \\ &= (I \otimes \alpha)(1 - \widehat{\nu \otimes P})^{-1}(T \otimes \beta) \end{aligned} \tag{31}$$

when $\rho(\widehat{\nu \otimes P}) < 1$, thanks to Lemma 2. This shows that the partition function Z is finite when the spectral radius $\rho(\widehat{\nu \otimes P})$ is strictly less than one, and that Z is a rational function of the entries of α, P, β, since the coefficients of the inverse of a matrix are rational in the entries of the matrix. $\square$

Example 3. Let $\Sigma = \{a, b\}$ and $\mathscr{F} = \{ab^2\}$. Using for instance the algorithm of derivatives [16, Chap. 5, Th. 2], we obtain the deterministic automaton depicted in Fig. 4, recognizing $\mathscr{S}_r = \Sigma^* ab^2 - \Sigma^* ab^2 \Sigma^+$. The states of this automaton are the nonempty languages $w^{-1}\mathscr{S}_r$, where $w \in \Sigma^*$, and $w^{-1}L = \{u \in \Sigma^* \mid wu \in L\}$. There is an arc from $w^{-1}\mathscr{S}_r$ to $a^{-1}w^{-1}\mathscr{S}_r$ with label a if $a^{-1}w^{-1}\mathscr{S}_r \neq \emptyset$. The initial node is $1^{-1}\mathscr{S}_r = \mathscr{S}_r$, the final nodes are such that $1 \in w^{-1}\mathscr{S}_r$. Here, $(ab^2)^{-1}\mathscr{S}_r$ is the unique final node. Since $\mathscr{S}_t$ is composed of the words with are prefixes of words in $\mathscr{S}_r$ but do not belong to $\mathscr{S}_r$, to obtain an automaton recognizing $\mathscr{S}$, we just have to mark all the states as final in Fig. 4. Thus, the characteristic function char $\mathscr{S}$ admits the linear representation (I, ν, F) such that:

$$p(a)\nu(a) + p(b)\nu(b) = \begin{bmatrix} p(b) & p(a) & \cdot & \cdot \\ \cdot & p(a) & p(b) & \cdot \\ \cdot & p(a) & \cdot & p(b) \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & \cdot & \cdot & \cdot \end{bmatrix}, \quad T = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

(the 0 entries are represented by dots). These matrices can be read directly on Fig. 4. For instance, we have $\nu(a)_{12} = 1$ since there is one arc from node 1 to node 2 with label a in the automaton, and $I_1 = 1$ since 1 is the initial state. Assuming, for simplicity, that p is a Bernoulli measure, we get the following rational expression for Z:

$$Z = I(1 - p(a)\nu(a) - p(b)\nu(b))^{-1}T = \frac{1 + p(a)\,(p(b))^2}{(1 - p(a) - p(a)p(b))(1 - p(b))} \ .$$

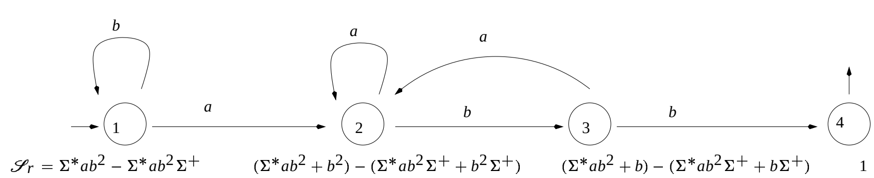


Figure 4: An automaton that recognizes $\Sigma^* ab^2 - \Sigma^* ab^2 \Sigma^+$.

Remark 3. There is a backward/forward duality between the automaton in Example 3 and the definition of the Lyapunov exponents: in (6), words are read from right to left, but to recognize $\mathscr{S}_r$ and $\mathscr{S}_t$, we use automata that read words from left to right, as usual.

4 Series Expansions of Lyapunov Exponents

4.1 Case of Stationary Probability Measures

All the results of this section need the following key assumption.

Memory Loss Property. *There is a subset $\mathscr{F} \subset \Sigma^+$, such that for all $z \in \mathscr{F}$, the image of the map $y \mapsto zy, \mathbb{R}^d \to \mathbb{R}^d$, is a line.*

For the heap model of Fig. 1, the memory loss property is satisfied, with $\mathscr{F} = \{c\}$. For more general heaps models, the memory loss property holds, with a more complex set $\mathscr{F}$, provided that the set of pieces cannot be split in independent subsets [24, 12, 19].

Our main theorem is the following.

Theorem 2 (Series Expansion Formula). *For all stationary probability measures p on Σ^*, such that*

$$\sum_{w\in\mathscr{S}} p(w)|w| < +\infty \ , \tag{32}$$

the Lyapunov exponent is given by

$$\Gamma = \lim_n D_n = \sum_{w\in\mathscr{S}_r} D(w) \ . \tag{33}$$

Proof. Since Γ is by definition the Cesaro limit of D_n, and since a convergent sequence Cesaro converges to the same limit, it suffices to prove the second equality in (33). Partitioning the sum in (13) in sums over equivalence classes modulo $\equiv_{\mathscr{F}}$ (see Prop. 2), we have:

$$\begin{aligned} D_n &= \sum_{w\in\Sigma^{n-1}} D(w) = \sum_{z\in\mathscr{S}} \sum_{\substack{w\equiv_{\mathscr{F}} z\\ |w|=n-1}} D(w) \\ &= \underbrace{\sum_{\substack{z\in\mathscr{S}_t\\ |z|=n-1}} D(z)}_{A_n} + \underbrace{\sum_{\substack{z\in\mathscr{S}_r\\ |z|\le n-1}} \left(\sum_{|u|=n-1-|z|} D(zu) \right)}_{B_n} \ . \end{aligned} \tag{34}$$

To compute the limit of D_n, we need an a priori estimate on $D(w)$. It will be useful to write $D(w) = \sum_{a\in\Sigma} p(aw)\psi(a, w)$, with

$$\psi(a, w) = \varphi awx - \varphi wx \ .$$

Using the non-expansiveness and homogeneity of the maps $x \mapsto bx$, with $b \in \Sigma$, it is immediate to show (by induction on the length of w) that

$$||\varphi wx|| \le (|w|K + K') + \|x\|$$

where $K = \max_{b\in\Sigma} \|b0\|$, and $K' = \|\varphi 0\|$. Hence,

$$|D(w)| \le \sum_{a\in\Sigma} p(aw)|\psi(a, w)| \le p(w)((2|w| + 1)K + 2K' + 2\|x\|) \ . \tag{35}$$

If the series (32) is convergent, then, evidently, the first term in the sum (34), A_n, tends to zero when $n \to \infty$. Let us now compute B_n. We have

$$\begin{aligned}
B_n &= \sum_{\substack{z\in\mathscr{S}_r\\ |z|\leq n-1}} \left(\sum_{|u|=n-1-|z|} \sum_{a\in\Sigma} p(azu)\psi(a,zu) \right) \\
&= \sum_{\substack{z\in\mathscr{S}_r\\ |z|\leq n-1}} \left(\sum_{|u|=n-1-|z|} \sum_{a\in\Sigma} p(azu)\psi(a,z) \right) \quad \text{(by (25))} \\
&= \sum_{\substack{z\in\mathscr{S}_r\\ |z|\leq n-1}} \sum_{a\in\Sigma} p(az)\psi(a,z) \quad \text{(by stationarity of } p) \\
&= \sum_{\substack{z\in\mathscr{S}_r\\ |z|\leq n-1}} D(w) \ ,
\end{aligned}$$

which shows (33). □

We next give a more explicit series expansion of Γ when p is rational.

Corollary 1. *If p is a stationary rational probability measure with linear representation (α, P, β), we have, as soon as* (32) *is satisfied:*

$$\Gamma = \sum_{w\in\mathscr{S}_r} \alpha\delta(wx)P(w)\beta \ . \tag{36}$$

In particular, in the Bernoulli case:

$$\Gamma = \sum_{w\in\mathscr{S}_r} \delta(wx)p(w) \ . \tag{37}$$

As a second corollary of Theorem 2, we obtain an estimate of the analyticity domain of the Lyapunov exponent, for Bernoulli probability measures. In this case, we will identify p with the vector $(p(a))_{a\in\Sigma} \in \mathbb{C}^\Sigma$, and we denote by $|p|$ the vector with entries $(|p(a)|)_{a\in\Sigma}$.

Corollary 2. *If p is a Bernoulli probability measure, the Lyapunov exponent Γ can be extended to an analytic function on the domain:*

$$\mathscr{D} = \{p \in \mathbb{C}^\Sigma \mid Z(|p|) < +\infty\} \ .$$

Proof. Formula (33) yields a representation of Γ as a power series in the variables $p(a)$, for $a \in \Sigma$, and we have seen in the proof of Theorem 2 that this power series is converging when $\sum_{w\in\mathscr{S}} |p(w)||w| < +\infty$. An elementary result of complex analysis ([13, Chap. IV, Prop. 3.2]) shows that for all $a \in \Sigma$, the partial derivative $\partial_{p(a)}Z(p)$ exists when $p \in \mathscr{D}$, and that it is given by the absolutely convergent series: $\partial_{p(a)}Z(p) = \sum_{z\in\mathscr{S}} p(a)^{-1}p(z)|z|_a$, where $|z|_a$ denotes the number of occurrences of the letter a in z. Hence, the sum $\sum_{z\in\mathscr{S}} p(z)|z| = \sum_{a\in\Sigma} p(a)\partial_{p(a)}Z(p)$ is absolutely convergent in $\mathscr{D}$, which shows that Γ is analytic in $\mathscr{D}$. □

When $\mathscr{F}$ is rational, Prop. 4 shows that Z is rational, and Corollary 2 yields an effective estimate: the power series in (33) converges on any polydisc centered at 0 that does not contain a pole of Z.

To extend Corollary 2 to the case of rational probability measures, we introduce an additional notation: if (α', P', β') is a r-dimensional complex valued linear representation, we denote by $(|\alpha'|, |P'|, |\beta'|)$ the nonnegative linear representation defined by $|\alpha'|_i = |\alpha'_i|$, $|P'|(a)_{ij} = |P'(a)_{ij}|$, $|\beta'|_i = |\beta'_i|$ (we warn the reader that $|P'(w)| \neq |P'|(w)$ in general).

Corollary 3. *If p is a stationary rational probability measure with linear representation (α, P, β), the Lyapunov exponent Γ can be extended to an analytic function of the entries of (α, P, β) on the domain:*

$$\begin{aligned} \mathscr{D} &= \{(\alpha, P, \beta) \mid \sum_{w \in \mathscr{S}_r} |\alpha||P|(w)|\beta| < +\infty\} \\ &\supset \{(\alpha, P, \beta) \mid \rho(\widehat{\nu \otimes |P|}) < 1\} \ . \end{aligned} \tag{38}$$

Proof. The proof is identical to that of Corollary 2, the only new ingredient being the inclusion in (38), which is provided by Prop. 4. □

Remark 4. Formula (37) is a special case of Furstenberg's cocycle formula (21). Indeed, we get a cocycle representation of $w \mapsto \varphi w x$ by taking: $S = \mathbb{F}(\Sigma, \mathscr{F})$, equipped with the action $\Sigma^* \times S \to S$ $(u, \overline{v}) \mapsto \overline{uv}$ and the cocycle $\kappa(u, \overline{v}) = \varphi u v x - \varphi v x$ (these quantities are independent of the representative v of $\overline{v}$); the equivalence class of the unit word as initial state; and the constant function $\varpi(\overline{v}) = \varphi x$ as output map. Prop. 3 shows that the underlying Markov chain on S, which coincides with the left random walk on $\mathbb{F}(\Sigma, \mathscr{F})$, has the unique invariant measure $\pi(\overline{w}) = p(w)$ for all $w \in \mathscr{S}_r$. Then, (37) coincides with (21).

We conclude this section by mentioning two consistency properties. The first one shows that the larger the set of forgetful factor is, the better the estimation of the analyticity domain of Γ is (for simplicity, we only consider the Bernoulli case).

Proposition 5. *If $\mathscr{F} \subset \mathscr{F}'$, and if p is Bernoulli, then $\mathscr{D}(\mathscr{F}) \subset \mathscr{D}(\mathscr{F}')$.*

Proof. Since $\mathscr{S}_r \subset \mathscr{S}_t \Sigma$, $\sum_{w \in \mathscr{S}_t} |p(w)| < +\infty \implies \sum_{w \in \mathscr{S}_r} |p(w)| \leq \sum_{z \in \mathscr{S}_t, a \in \Sigma} |p(za)| \leq (\sum_{z \in \mathscr{S}_t} |p(z)|)(\sum_{a \in \Sigma} |p(a)|) < +\infty$. Hence, $Z(|p|) < +\infty$ if, and only if, $\sum_{w \in \mathscr{S}_t} |p(w)| < \infty$, Since $\mathscr{S}_t(\mathscr{F}) = \Sigma^* - \Sigma^* \mathscr{F} \Sigma^*$, $\mathscr{F} \subset \mathscr{F}' \implies \mathscr{S}_t(\mathscr{F}) \supset \mathscr{S}_t(\mathscr{F}')$, which gives $\mathscr{D}(\mathscr{F}) \subset \mathscr{D}(\mathscr{F}')$. □

The following corollary shows that our analyticity domains contain the set of "real" probabilities, perhaps up to boundaries.

Corollary 4. *If $\mathscr{F} \neq \emptyset$, then, $\mathscr{D}(\mathscr{F})$ contains the set of positive Bernoulli probability measures.*

Proof. By Prop. 5, it is enough to check this when $\mathscr{F}$ is reduced to a single word: this is an elementary exercise of calculus that we leave to the reader. □

4.2 Case of Nonstationary Rational Probability Measures

To extend Theorem 2 to the case of non stationary rational probability measures, we recall some properties of eigenprojectors. The *eigenprojector* for an eigenvalue λ of a matrix A is defined by

$$\Pi = \frac{1}{2i\pi} \int_\gamma (z - A)^{-1} dz \ , \tag{39}$$

where the integral is taken over a circle γ containing only the eigenvalue λ (see [33, Chap. II,§ 1.4]). We say that λ is *semisimple* if there is no nilpotent term in the Jordan decomposition of A for the eigenvalue λ, or equivalently, if λ is a simple pole of $(z - A)^{-1}$ (see [33, Chap. I,§ 5.4]). Then,

$$\Pi = \lim_{z \to \lambda} (z - \lambda)(z - A)^{-1} \ . \tag{40}$$

Given a (nonnegative) linear representation (α, P, β) of dimension r of a rational probability measure p, we say that the index $i \in \{1, \dots, r\}$ is *accessible* (resp. *co-accessible*) if there exists $w \in \Sigma^*$ such that $(\alpha P(w))_i \neq 0$ (resp. $(P(w)\beta)_i \neq 0$). We say that a linear representation is *trim* if all the $i \in \{1, \dots, r\}$ are both accessible and co-accessible. Clearly, if i is not accessible, or not co-accessible, the linear representation obtained by deleting column i of α and P, and row i of P and β, is still a linear representation of p. Thus, there is no loss of generality in considering only trim representations. We remark that a (nonnegative) linear representation is trim iff the following holds:

$$\forall i, j, \quad \exists k, l, \quad (\alpha \hat{P}^k)_i \neq 0, \text{ and } (\hat{P}^l \beta)_j \neq 0 \quad . \tag{41}$$

We shall use the following elementary observation. Recall that the *Perron root* of a nonnegative matrix is by definition its spectral radius, which is an eigenvalue associated to a nonnegative eigenvector, by the Perron-Frobenius theorem [6].

Lemma 3. *If (α, P, β) is a (nonnegative) trim linear representation of a rational probability measure p, then, the Perron root of the matrix $\hat{P}$ is equal to 1 and is semisimple. Besides, the eigenprojector of $\hat{P}$ for the eigenvalue 1 is equal to:*

$$\Pi = \lim_{N \to \infty} \frac{\sum_{k=1}^{N} \hat{P}^k}{N} \tag{42}$$

Proof. Since p is a probability measure, we have:

$$1 = \sum_{w \in \Sigma^n} p(w) = \alpha \sum_{w \in \Sigma^n} p(w)\beta = \alpha \hat{P}^n \beta \quad . \tag{43}$$

Using (41), we easily derive from (43) the existence of $K > 0$ such that $\hat{P}^n_{ij} \leq K$, for all i, j and n. Then,

$$\rho(\hat{P}) = \lim_{n \to \infty} \|\hat{P}^n\|^{\frac{1}{n}} = \lim_{n \to \infty} \sup_{ij} (\hat{P}^n_{ij})^{1/n} \leq 1 \quad .$$

But $\rho(\hat{P}) < 1$ would imply that $1 = \alpha \hat{P}^n \beta \to 0$, a contradiction, which shows that $\rho(\hat{P}) = 1$. Moreover, since $\hat{P}^n_{ij} \leq K$, the entries of the matrix $(1-q)(1-q\hat{P})^{-1}$ remain bounded by $(1-q)\sum_{k \geq 0} q^k K = K$ when $q \to 1^-$. Thus, $(\hat{P} - z)^{-1}$ has only a pole of order 1 at $z = 1$, which means that 1 is semisimple. Then, the convergence of $\frac{\sum_{k=0}^{N} \hat{P}^k}{N}$ is a consequence of (40), together with the Tauberian theorem of Hardy and Littlewood, already used in the proof of Theorem 1. ☐

We are now in position to state the analogue of Theorem 2 for rational nonstationary probability measures.

Theorem 3. *If (α, P, β) is a (nonnegative) trim linear representation of a rational probability measure p such that $\rho(\widehat{\nu \otimes P}) < 1$, then, the Lyapunov exponent exists, and it is equal to:*

$$\Gamma = \sum_{w \in \mathscr{S}_r} \alpha \delta(wx) P(w) \Pi \beta \quad , \tag{44}$$

where Π denotes the eigenprojector of $\hat{P}$ for the eigenvalue 1.

Proof. In the non-stationary case, $\lim_n D_n$ does not exist, in general, but we shall prove that $\lim_n n^{-1}S_n = \lim_n n^{-1}(D_1 + \cdots + D_n)$ does exist, when $\sum_{w\in\mathscr{S}} p(w) < +\infty$. Using the decomposition in (34), we can write:

$$S_n = A'_n + B'_n, \text{ where } A'_n = A_1 + \cdots + A_n, \quad B'_n = B_1 + \cdots + B_n \ .$$

The bound (35) shows that A_n tends to zero, hence, A'_n/n tends to zero. For B'_n, we can write, from (34),

$$\begin{aligned} B'_n &= \sum_{\substack{z\in\mathscr{S}_r \\ |z|\leq n-1}} \sum_{|u|\leq n-1-|z|} \sum_{a\in\Sigma} \alpha P(azu)\beta\psi(a, zu) \\ &= \sum_{\substack{z\in\mathscr{S}_r \\ |z|\leq n-1}} \sum_{|u|\leq n-1-|z|} \sum_{a\in\Sigma} \alpha P(az)P(u)\beta\psi(a, z) \\ &= \sum_{\substack{z\in\mathscr{S}_r \\ |z|\leq n-1}} \sum_{a\in\Sigma} \alpha P(az)\psi(a, z)(\sum_{|u|\leq n-1-|z|} P(u))\beta \\ &= \sum_{\substack{z\in\mathscr{S}_r \\ |z|\leq n-1}} \sum_{a\in\Sigma} \alpha P(az)\psi(a, z)(1 + \hat{P} + \cdots + \hat{P}^{n-1-|z|})\beta \ . \end{aligned}$$

If $\rho(\widehat{\nu\otimes P}) < 1$, then, using Lemma 2 and arguing as in the proof of Corollary 2, we get that the (nonnegative) power series $\sum_{w\in\mathscr{S}} P(w)|w|$ is convergent. Since $n^{-1}(1 + \hat{P} + \cdots + \hat{P}^{n-1-|z|}) \to \Pi$, and since $\sum_{a\in\Sigma} P(az)\psi(a, z) = \delta(z)$, by a dominated convergence argument, $n^{-1}B'_n \to \sum_{z\in\mathscr{S}_r} \alpha\delta(zx)P(z)\Pi\beta$. □

Remark 5. Theorem 3 implies in particular that the Lyapunov exponent (7) exists under the memory loss assumption, even if the probability measure p is *not* stationary, provided that it is rational.

Remark 6. Theorem 3 yields the following local analyticity result. Let $\kappa = (\kappa_1, \ldots, \kappa_m) \in \mathbb{C}^m$ and let $(\alpha_\kappa, P_\kappa, \beta_\kappa)$ denote a linear representation whose entries are analytic functions of the κ_i, near $\kappa = 0$, with α_0, P_0, and β_0 nonnegative, and $\rho(\widehat{\nu\otimes P_0}) < 1$. Let us assume that 1 is a simple eigenvalue of P_0. Then, by classical results of perturbation theory [33, Chap. II,§ 1.4], there is a neighborhood V of 0 such that there is a unique eigenvalue λ_κ of $\hat{P}_\kappa$, depending analytically of κ, such that $\lambda_0 = 1$. The spectral projector for λ_κ, Π_κ, is analytic on V. Moreover, by continuity of the Perron root, possibly after restricting V, we may also assume that $\rho(\widehat{\nu\otimes |P_\kappa|}) < 1$ for all $\kappa \in V$, so that the sum (44) extends Γ analytically to V. (This result holds more generally when 1 is a semisimple eigenvalue of $\hat{P}$, provided that the multiplicity of λ_κ remains constant in V.)

5 Applications

5.1 The Case $\mathscr{F} = \{a^c\}$ under a Bernoulli Measure

The case when $\mathscr{F} = \{a^c\}$, for some letter $a \in \Sigma$, has been considered in detail in [3]. The importance of this case stems from the max-plus spectral theorem which shows that if a is a max-plus linear map, rather generically — if the matrix of a is irreducible and if its critical graph has a single strongly connected component with cyclicity 1 — there is a power of a whose image is a line (the max-plus spectral theorem has been proved by various authors, see [1, 39, 27, 5] for recent references). When $\mathscr{F} = \{a^c\}$, it is easy to see that

$$\mathscr{S}_r = (\{a, \ldots, a^{c-1}\}B^+)^* a^c \cup (B^+\{a, \ldots, a^{c-1}\})^* B^+ a^c \ , \tag{45}$$

where $B = \Sigma \setminus \{a\}$, and where, as usual, we write a^c instead of $\{a^c\}$. Since the rational expression in (45) is unambiguous [35], we obtain $Z_r \stackrel{\text{def}}{=} \sum_{w \in \mathscr{S}_r} p(w)$ by replacing $\cup$ by $+$, concatenation by product, and $(\cdot)^*$ by $(1-\cdot)^{-1}$ in (45). After simplification:

$$Z_r = \frac{p(a)^c}{1 - (\sum_{b \in \Sigma \setminus \{a\}} p(b))(1 + p(a) + \cdots + p(a)^{c-1})} \quad . \tag{46}$$

The argument of the proof of Prop. 5 shows that Z and Z_r have the same absolute convergence domain. Thus, by specialization of Theorem 2, we obtain the following simple series expansion (compare with [3]), which allows us to get the optimal estimate of the convergence radius.

Corollary 5. *Assume that $\mathscr{F} = \{a^c\}$ for some $a \in \Sigma$, and $c > 0$. (i) The Lyapunov exponent is given by*

$$\Gamma = \sum_{w \in \mathscr{S}_r} \delta(wx) p(w) = \lim_{n \to \infty} D_n \quad , \tag{47}$$

for all Bernoulli measures p such that $Z(p) < +\infty$, and Γ can be extended analytically to the domain

$$(\sum_{b \in \Sigma \setminus \{a\}} |p(b)|)(1 + |p(a)| + \cdots + |p(a)|^{c-1}) < 1 \quad . \tag{48}$$

(ii) In particular, when $\Sigma = \{a, b\}$, the Taylor series of Γ, seen as a function of $p(b)$, has convergence radius $2^{\frac{1}{c}} - 1$, and this bound is tight.

Proof. Part *(i)* of the theorem readily follows from Theorem 2, together with (46). To prove part *(ii)*, we write as $1 - F(p)$ the denominator of (46). The Taylor series of Γ converges absolutely at any $p(b)$ such that $(p(a) = 1 - p(b), p(b))$ lies in the domain $F(|p|) < 1$. Since $F(|p|) = |p(b)|(1 + |1 - p(b)| + \cdots + |1 - p(b)|^{c-1}) \leq |p(b)|(1 + 1 + |p(b)| + \cdots + (1 + |p(b)|)^{c-1})$, a sufficient condition for $F(|p|) < 1$ is

$$|p(b)| \left(\frac{(1 + |p(b)|)^c - 1}{|p(b)|} \right) = (1 + |p(b)|)^c - 1 < 1 \quad ,$$

which is the case if $|p(b)| < 2^{\frac{1}{c}} - 1 = \sqrt{2} - 1$. When $c = 2$, the optimality of the bound $2^{\frac{1}{c}} - 1$ is clear from Example 2, since in this case, Γ has a pole at $p(b) = 1 - \sqrt{2}$ (see Formula (22)). The optimality of the bound, for a general c, is shown by Example 4 below. $\square$

Example 2 shows that substituting $p(a) = 1 - p(b)$ in the power series (33), or more generally, looking for Taylor series expansion, is not always a good thing to do: here the series (33), seen as a series in $p(a), p(b)$, converges when $|p(b)|(1 + |p(a)| + \cdots + |p(a)|^{c-1}) < 1$, it is easy to sum, and its convergence domain contains the physically interesting domain $p(b) \in [0, 1[$, $p(a) = 1 - p(b)$, whereas the Taylor series of Γ, seen as a function of $p(b)$, which is obtained by substitution of $p(a) = 1 - p(b)$, is divergent at $p(b) = 2^{\frac{1}{c}} - 1$.

Example 4. For a given $c \geq 2$, we consider the max-plus linear maps a, b associated to the $(c+1) \times (c+1)$ matrices

$$A = \begin{pmatrix} 1 & 1 & -\infty & \cdots & -\infty & 0 \\ \vdots & -\infty & \ddots & \ddots & \vdots & -\infty \\ \vdots & \vdots & \ddots & 1 & -\infty & \vdots \\ 1 & \vdots & & \ddots & 0 & \vdots \\ 2 & -\infty & \cdots & \cdots & -\infty & -\infty \\ 2 & -\infty & \cdots & \cdots & -\infty & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & \cdots & \cdots & 1 & 0 & 0 \\ \vdots & & & \vdots & \vdots & \vdots \\ \vdots & & & \vdots & \vdots & \vdots \\ 1 & \cdots & \cdots & 1 & 0 & 0 \\ 2 & \cdots & \cdots & 2 & 0 & c+1 \\ 2 & \cdots & \cdots & 2 & 1 & 0 \end{pmatrix}.$$

The underlying Markov chain, built as in Example 2, has the $c+1$ states: $(0,\ldots,0,1,1)^T$, $(0,\ldots,0,c+1,1)^T$, $(0,\ldots,0,c,1,1)^T$, $(0,\ldots,0,c,0,1,1)^T$, $\ldots$,$(0,0,c,0,\ldots,0,1,1)^T$, $(0,c,0,\ldots,0,1,1)^T$, $(c,0,\ldots,0,1,1)^T$; its transition matrix is

$$M = \begin{pmatrix} 1-p(b) & p(b) & 0 & \cdots & 0 \\ p(b) & 0 & 1-p(b) & \ddots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ p(b) & \vdots & \vdots & \ddots & 1-p(b) \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{49}$$

it has the invariant measure

$$m = \frac{1}{2-(1-p(b))^c}\begin{pmatrix} 1 & p(b) & p(b)(1-p(b)) & \ldots & p(b)(1-p(b))^{c-1} \end{pmatrix},$$

and the associated cocycle κ is determined by the following vectors of dimension $c+1$:

$$\kappa(a,\cdot) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ c+1 \end{pmatrix}, \qquad \kappa(b,\cdot) = \begin{pmatrix} 1 \\ c+1 \\ \vdots \\ c+1 \\ c+1 \end{pmatrix}.$$

Using Formula (21), we get

$$\Gamma = m[(1-p(b))\kappa(a,\cdot) + p(b)\kappa(b,\cdot)] = \frac{1+cp(b)-(1-p(b))^c}{2-(1-p(b))^c},$$

which has a pole at $1-2^{\frac{1}{c}}$.

5.2 Random Heaps of Pieces

To illustrate the representation formula (37), we generalize the heap model of Fig. 1, by taking $d+1$ pieces with associated operators

$$\begin{aligned} a_0 x &= (1+\max_{1\le k\le d} x_k,\ldots,1+\max_{1\le k\le d} x_k)^T, \\ a_j x &= (x_1,\ldots,x_{j-1},1+x_j,x_{j+1},\ldots,x_d)^T, \quad \forall 1\le j\le d \end{aligned}$$

(for $j>0$, the piece corresponding to a_j occupies column j, the piece corresponding to a_0 occupies all the columns, and all the pieces have height 1). We take $\varphi = \mathsf{t}$. Here, $\mathscr{F} = \{a_0\}$, and,

$$\delta(y) = \sum_{1\le i\le d} p(a_i)(\mathsf{t}a_i y - \mathsf{t}y) = \sum_{i\in\arg\max_j y_j} p(a_i)\ .$$

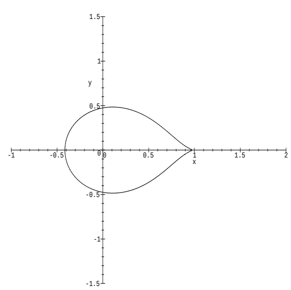
For all $I \subset \{1, \ldots, d\}$, we set $L^I = \{w \in \{a_1, \ldots, a_d\}^* \mid \forall i, j \in I, \forall k \notin I,\ |w|_i = |w|_j > |w_k|\} = \{w \in \{a_1, \ldots, a_d\}^* \mid \arg\max_j (w0)_j = I\}$, and $Z^I = \sum_{w \in L^I} p(w)$. We have, by direct application of Theorem 2, the following representation, which was obtained by independent means by Krob [34].

Proposition 6. *The Lyapunov exponent of the above heap model is*

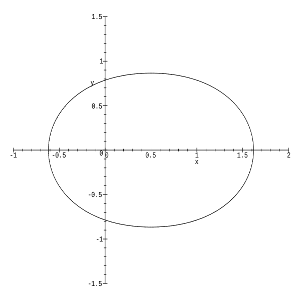
$$\Gamma = \sum_{I \subset \{1,\ldots,d\}} Z^I \Big(\sum_{i \in I \cup \{0\}} p(a_i) \Big) \ . \tag{50}$$

5.3 Multiple Memory Loss Relations

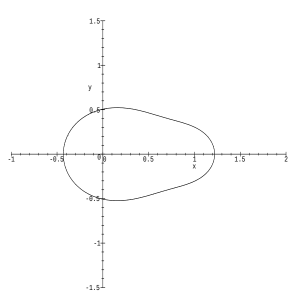
Finally, we show the absolute convergence domains $Z(|p|) < +\infty$ associated to various finite sets $\mathscr{F}$ (by Corollary 2, the Lyapunov exponent is analytic on these domains). We set $z = p(b)$, and we represent $\Re[z]$ and $\Im[z]$ on the x and y axes, respectively, so that the segment $y = 0$ and $x \in [0, 1]$ represents the "real" probability region. The domains are obtained from Prop. 4, in the case of Bernoulli measures over the alphabet $\Sigma = \{a, b\}$. When $\mathscr{F} = \{a^2\}$, the domain was already given in Corollary 5. The case when $\mathscr{F} = \{ab^2\}$ was considered in Example 3. For each of the other cases, it is easy to write an unambiguous rational expression for $\mathscr{S}_t$ or $\mathscr{S}_r$, from which the absolute convergence domain of Z can be obtained. We checked these computations using AMoRE [40] and MAPLE.



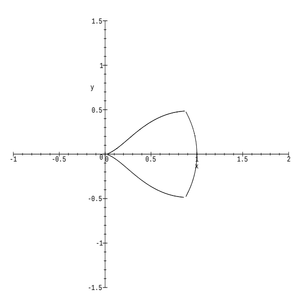
$\mathscr{F} = \{a^2\}$
$|z|(1 + |1 - z|) < 1$



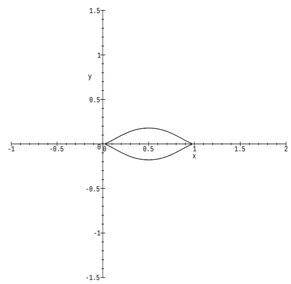
$\mathscr{F} = \{a^2, b^2\}$
$|1 - z||z| < 1$



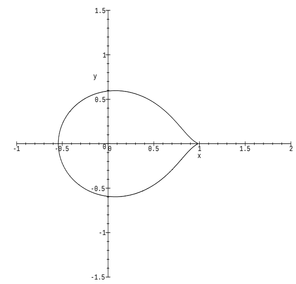
$\mathscr{F} = \{a^2, b^4\}$
$|1 - z|(|z| + |z|^2 + |z|^3) < 1$



$\mathscr{F} = \{ab^2\}$
$|1 - z|(1 + |z|) < 1$ and $|z| < 1$



$$\mathscr{F} = \{(ab)^2\}$$
$$-|z|^2|1-z|^2 + |z||1-z|^2 +$$
$$+|z|^2|1-z| - |z||1-z| + |1-z| + |z| < 1$$



$$\mathscr{F} = \{(ab)^2, a^2\}$$
$$|z|(1 + |1-z||z|) < 1$$

Acknowledgement. The authors thank François Baccelli for many useful discussions.

References

[1] F. Baccelli, G. Cohen, G.J. Olsder, and J.P. Quadrat. *Synchronization and Linearity*. Wiley, 1992.

[2] F. Baccelli, S. Gaubert, and D. Hong. Representation and expansion of max-plus Lyapunov exponents. In *Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, 1999.

[3] F. Baccelli and D. Hong. Analytic expansions of max-plus Lyapunov exponents. To appear in Annals of Appl. Prob.

[4] F. Baccelli and D. Hong. Analyticity of iterates of random non-expansive maps. *Adv. Appl. Prob.*, 32(1). To appear in the issue of March 2000.

[5] R. B. Bapat. A max version of the Perron-Frobenius theorem. In *Proceedings of the Sixth Conference of the International Linear Algebra Society (Chemnitz, 1996)*, volume 275/276, pages 3–18, 1998.

[6] A. Berman and R.J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Academic Press, 1979.

[7] J. Berstel. *Transductions and Context-Free languages*. Teubner, Stuttgart, 1979.

[8] J. Berstel and D. Perrin. *Theory of codes*. Academic Press Inc., Orlando, Fla., 1985.

[9] J. Berstel and C. Reutenauer. *Rational Series and their Languages*. Springer, 1988.

[10] V. D. Blondel, S. Gaubert, and J. N. Tsitsiklis. Approximating the spectral radius of sets of matrices in the max-algebra is NP-hard. May 1999. To appear in the IEEE Trans. on Automatic Control.

[11] P. Bougerol and J. Lacroix. *Products of Random Matrices with Applications to Schrödinger operators*, volume 8. Birkhaüser, 1985.

[12] M. Brilman and J.M. Vincent. Dynamics of synchronized parallel systems. *Comm. Statist. Stochastic Models*, 13(3):605–617, 1997.

[13] H. Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Hermann, 1985.

[14] J. E. Cohen. Erratum: "Subadditivity, generalized products of random matrices and operations research". *SIAM Rev.*, 35(1):124, 1993.

[15] J.E. Cohen. Subadditivity, generalized products of random matrices and Operations Research. *SIAM Rev.*, 30(1):69–86, 1988. Erratum in [14].

[16] J.H. Conway. *Regular algebra and finite machines*. Chapman and Hall, 1971.

[17] J. Cook and B. Derrida. Lyapunov exponents of large, sparse random matrices and the problem of directed polymers with random complex weights. *J. Stat. Phys.*, 61:961–986, 1990.

[18] M.G. Crandall and L. Tartar. Some relations between non expansive and order preserving maps. *Proceedings of the AMS*, 78(3):385–390, 1980.

[19] V. Dumas and Ph. Robert. On the maximal throughput of a synchronization process. Submitted to Mathematics of Operation research, 1998.

[20] Simon P. Eveson and Roger D. Nussbaum. An elementary proof of the Birkhoff-Hopf theorem. *Math. Proc. Cambridge Philos. Soc.*, 117(1):31–55, 1995.

[21] H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963.

[22] E. Gardner and B. Derrida. Zero temperature magnetization of a one-dimensional spin glass. *J. Stat. Phys.*, 39:367–377, 1985.

[23] S. Gaubert. Performance evaluation of (max,+) automata. *IEEE Trans. on Automatic Control*, 40(12), Dec 1995.

[24] S. Gaubert and J. Mairesse. Task resource systems and (max,+) automata. In J. Gunawardena, editor, *Idempotency*. Cambridge University Press, 1995. 1998.

[25] S. Gaubert and J. Mairesse. Asymptotic analysis of heaps of pieces and application to timed Petri nets. In *PNPM'99*, Saragoza, Spain, Sep. 1999.

[26] S. Gaubert and J. Mairesse. Modeling and analysis of timed Petri nets using heaps of pieces. *IEEE Trans. Automat. Control*, 44(4):683–697, 1999.

[27] S. Gaubert and M. Plus. Methods and applications of (max,+) linear algebra. In R. Reischuk and M. Morvan, editors, *STACS'97*, number 1200 in LNCS, Lübeck, March 1997. Springer.

[28] B. Gaujal and A. Jean Marie. Computational issues in stochastic recursive systems. In J. Gunawardena, editor, *Idempotency*. Cambridge University Press, 1998.

[29] J. Gunawardena and M. Keane. On the existence of cycle times for some nonexpansive maps. Technical Report HPL-BRIMS-95-003, Hewlett-Packard Labs, 1995.

[30] G. Hansel and D. Perrin. Mesures de probabilités rationnelles. In M. Lothaire, editor, *Mots*. Hermes, 1990.

[31] G.H. Hardy. *Divergent Series*. Chelsea, 1991. Second, reprinted edition.

[32] B. Hennion. Limit theorems for products of positive random matrices. *The Annals of Probability*, 25(4):1545–1587, 1997.

[33] T. Kato. *Perturbation Theory for Linear Operators*. Springer, 1980.

[34] D. Krob. Personal communication. 1997.

[35] G. Lallement. *Semigroups and Combinatorial Applications*. Wiley, 1979.

[36] E. Le Page. Répartition d'état d'un opérateur de Schrödinger aléatoire. In *Probability measures on groups, VII*, number 1064 in LNM, pages 309–367, Oberwolfach, 1984. Springer.

[37] J. Mairesse. Products of irreducible random matrices in the (max,+) algebra. *Adv. Applied Prob.*, 29(2):444–477, June 1997.

[38] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities*. Allyn and Bacon, Boston, 1964.

[39] V. Maslov and S. Samborskiĭ, editors. *Idempotent analysis*, volume 13 of *Adv. in Sov. Math.* AMS, RI, 1992.

[40] O. Matz, A. Miller, A. Potthoff, W. Thomas, and E. Valkema. Report on the program AMoRE. Technical Report 9507, Christian-Alberchts-Universität, Kiel, October 1995. ftp.informatik.uni-kiel.de, directory: pub/kiel/amore/.

[41] G.J. Olsder, J.A.C. Resing, R.E. de Vries, M.S. Keane, and G. Hooghiemstra. Discrete event systems with stochastic processing times. *IEEE Trans. on Automatic Control*, 35:299–302, 1990.

[42] M. Pelletier and J. Sakarovitch. Easy multiplications ii. extentions of rational semigroups. *Information and Computation*, 88(1):18–59, September 1990.

[43] Y. Peres. Domains of analytic continuation for the top Lyapunov exponent. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(1):131–148, 1992.

[44] N. Saheb. Concurrency measure in commutation monoids. *Discrete Applied Mathematics*, 24:223–236, 1989.

[45] J. Sakarovitch. Easy multiplications i. the realm of Kleene's theorem. *Information and Computation*, 74(3):173–197, September 1987.

[46] C. Sparrow. Personal communication. 1998.

[47] J.M. Vincent. Some ergodic results on stochastic iterative discrete events systems. *DEDS: Theory and Applications*, 7(2):209–233, 1997.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399